PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82, 15/29, A01H 5/00, C12N 5/10

A1

(11) International Publication Number: WO 95/35386

(43) International Publication Date: 28 December 1995 (28.12.95)

(21) International Application Number: PCT/GB95/01443

(22) International Filing Date: 19 June 1995 (19.06.95)

(30) Priority Data:

| | | |
|---|---|---|
| 9412123.3 | 17 June 1994 (17.06.94) | GB |
| 9412124.1 | 17 June 1994 (17.06.94) | GB |

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(71) Applicant *(for all designated States except US)*: SCOTTISH CROP RESEARCH INSTITUTE [GB/GB]; Invergowrie, Dundee DD2 5DA (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BROWN, John, William, Slessor [GB/GB]; 49 Meadowview Drive, Inchture, Perthshire PH14 9TQ (GB). CLARK, Gillian, Patricia [GB/GB]; 7 Lowson Road, Forfar, Angus DD8 1LB (GB). SIMPSON, Gordon, Grant [GB/GB]; Friedrich, Miescher Institute, CH-4005 Basel (CH).

(74) Agent: PATTULLO, Norman; Murgitroyd & Company, 373 Scotland Street, Glasgow G5 8QA (GB).

(54) Title: EXPRESSION CONTROL POLYNUCLEOTIDES DERIVED FROM SPLICEOSOMAL PROTEIN GENE PROMOTERS

(57) Abstract

An expression control polynucleotide capable of affecting the expression of a second polynucleotide, and derived from a spliceosomal protein gene promoter. Isolation of plant spliceosomal protein gene promoters from potato and maize is described. Partial sequences of the promoters of two potato spliceosomal protein gene promoters are disclosed.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AT** | Austria | **GB** | United Kingdom | **MR** | Mauritania |
| **AU** | Australia | **GE** | Georgia | **MW** | Malawi |
| **BB** | Barbados | **GN** | Guinea | **NE** | Niger |
| **BE** | Belgium | **GR** | Greece | **NL** | Netherlands |
| **BF** | Burkina Faso | **HU** | Hungary | **NO** | Norway |
| **BG** | Bulgaria | **IE** | Ireland | **NZ** | New Zealand |
| **BJ** | Benin | **IT** | Italy | **PL** | Poland |
| **BR** | Brazil | **JP** | Japan | **PT** | Portugal |
| **BY** | Belarus | **KE** | Kenya | **RO** | Romania |
| **CA** | Canada | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CF** | Central African Republic | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CG** | Congo | | | **SE** | Sweden |
| **CH** | Switzerland | **KR** | Republic of Korea | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **KZ** | Kazakhstan | **SK** | Slovakia |
| **CM** | Cameroon | **LI** | Liechtenstein | **SN** | Senegal |
| **CN** | China | **LK** | Sri Lanka | **TD** | Chad |
| **CS** | Czechoslovakia | **LU** | Luxembourg | **TG** | Togo |
| **CZ** | Czech Republic | **LV** | Latvia | **TJ** | Tajikistan |
| **DE** | Germany | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **DK** | Denmark | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **US** | United States of America |
| **FI** | Finland | **ML** | Mali | **UZ** | Uzbekistan |
| **FR** | France | **MN** | Mongolia | **VN** | Viet Nam |
| **GA** | Gabon | | | | |

"Expression Control Polynucleotides derived from Spliceosomal Protein Gene Promoters"

This invention relates to expression control polynucleotides derived from spliceosomal protein gene promoters.

It is already well known that organisms such as plants or animals with novel characteristics can be produced by introducing genes or DNA sequences from the same or a related organism or by introducing genes or DNA sequences from other organisms. The key to production of novel phenotypes is the active expression of at least a part of the introduced DNA sequence. Generally expression of the introduced DNA sequence will occur only in the presence of an expression control polynucleotide, such as a promoter, which is compatible with the host organism. Promoters are nucleic acid sequences which are currently believed to regulate the expression of a gene by facilitating the binding of proteins required for transcription, such as RNA polymerase, to a portion of the nucleic acid sequence upstream of the gene. Thus for protein coding genes the DNA is transcribed into mRNA (messenger RNA) which is then relocated to the cytoplasm where it is

available for translation into polypeptide.

Whereas promoters are generally disposed upstream of the genes they regulate and are thought to act by providing a binding site for an RNA polymerase transcription complex, another form of expression control polynucleotide known as an "enhancer" sequence can control the expression of a gene usually without regard to its relative position or orientation.

Promoters can vary in levels of expression induction and in their expression patterns. Some promoters are active in a tissue-specific or developmental stage-specific manner while other promoters are active constantly to continually drive expression of the gene they control. Such constantly active promoters are called constitutive promoters. The promoters of this latter class most widely used in genetic engineering are the Cauliflower Mosaic Virus (CaMV) 35S RNA promoter, nopaline synthetase (nos) or octopine synthetase (ocs) promoters. Constitutive promoters induce gene expression at a relatively constant rate.

Although the above conventional promoters generally drive expression at high levels, many have the disadvantage that they are derived from plant infectious agents; the CaMV 35S RNA promoter is derived from a plant virus and the other promoters mentioned are derived from strains of *Agrobacterium*, soil-borne infectious bacteria. The source of these promoters is a cause of concern in transgenic plant production. In addition, detailed analysis of expression patterns of CaMV 35S promoter, have shown that its levels of expression can vary greatly among different plant tissues to the level where it is inactive in some tissues and is therefore, no longer constitutive for

such tissue. For many biotechnological objectives, constitutive expression in all cells and tissues would be of great advantage.

Splicesomal proteins are believed to be present in virtually all eukaryotic cells and are involved in the phenomenon of pre-mRNA splicing which removes introns (non-coding regions) from RNA transcripts before protein production.

The present invention seeks to provide a promoter (or other expression control polynucleotide) which is not derived from an infectious agent and which is suitable for use in the control of expression of recombinant genes in the construction of transgenic organisms such as plants and animals.

The present invention also seeks to provide a promoter (or other expression control polynucleotide) which is likely to be active throughout all or most cells of the organism.

According to the present invention there is provided an expression control polynucleotide at least partially derived from a spliceosomal protein gene promoter.

The expression control polynucleotide of the invention is capable of controlling the expression of a second polynucleotide (preferably comprising a polypeptide-encoding sequence) operably linked thereto.

RNA sequences which do not code for protein can also be expressed eg ribozymes or anti-sense RNA.

The term "expression control polynucleotide" as used herein will include promoters, enhancers or any other

functional equivalents or any other sequence elements which affect expression of other gene sequences.

The term "polypeptide-encoding sequence" as applied herein to polynucleotides means a polynucleotide comprising a sequence which can be transcribed into mRNA, which itself can be translated into a polypeptide. The "polypeptide-encoding sequence" may include non-translated portions, such as introns.

The spliceosomal protein gene promoter may be derived from plants. The plants may be dicotyledonous (eg potatoes), or monocotyledonous (eg maize).

The present invention also provides a recombinant polynucleotide comprising an expression control polynucleotide according to the invention operably linked to a second polynucleotide (preferably comprising a polypeptide-encoding sequence).

The present invention also provides a recombinant vector containing an expression control polynucleotide or a recombinant polynucleotide as defined above.

According to the present invention there is also provided a method of producing a recombinant vector, said method comprising ligating an expression control polynucleotide into a vector or part thereof. A method of producing a transformed cell by transfecting a host cell using said recombinant vector forms another aspect of the invention.

The present invention also provides a transformed host cell containing a recombinant polynucleotide or vector as defined above.

The present invention also provides a transgenic organism (for example a transgenic plant) containing a recombinant polynucleotide or vector as defined above. The progeny (and seeds) of such transgenic organisms forms a further part of the invention.

The present invention also provides a method for controlling the expression of a polypeptide from a nucleotide sequence encoding the polypeptide, said method comprising operably linking said sequence to an expression control polynucleotide of the invention.

The expression control polynucleotide of the invention may comprise double- or single-stranded DNA or RNA.

Three cultures of *E.coli* (SCRI/JB/1, SCRI/JB/2 and SCRI/JB/3), each containing a plasmid having an expression control polynucleotide according to the invention were deposited on 14 March 1994 with the National Collection of Type Cultures under numbers NCTC 12864, NCTC 12865 and NCTC 12866 respectively.

Cultures SCRI/JB/1 and 2 contain dicotyledonous spliceosomal protein gene expression control polynucleotides (promoters for potato U1A and U2B" genes respectively); SCRI/JB/3 contains a monocotyledonous expression control polynucleotide (promoter for a maize PRP8 gene).

Accordingly, the present invention also provides NCTC deposits Nos 12864, 12865 and 12866 and the plasmids thereof.

Recombinant DNA technology has been recognised as a powerful technique not only in research but also for commercial purposes. Thus, by using recombinant DNA

techniques (see Sambrook et al 1982 and "Principles of Genetic Engineering", Old and Primrose, 5th edition, 1994) exogenous genetic material can be transferred to a host cell and the polypeptide encoded by the exogenous genetic material may be replicated by and/or expressed within the host. For the purposes of simplicity recombinant DNA technology is normally carried out with prokaryotic micro-organisms, for example bacteria such as *E. coli*, as host. However, use has also been made of eukaryotic organisms, in particular yeasts or algae, and in certain applications eukaryotic cell cultures may also be used.

Genetic alterations to mammalian species by micro-injection of genes into the pro-nuclei of single-cell embryos is also well known and has been described by Brinster et al, in Cell 27: 223-231, 1981. Thus general techniques used in recombinant DNA technology and the production of transgenic organisms is within the scope of the skilled man.

Using a cDNA sequence (complimentary DNA - reverse transcribed from mRNA), either full-length or partial, as a probe for a gene of interest, the gene promoter can be readily isolated by standard procedures. Briefly, the cDNA probe is first used to screen a genomic library to isolate a genomic clone containing the promoter and coding sequence. Restriction mapping and Southern blotting with the cDNA as probe delineates the region of the genomic clone containing the coding sequence. Sequencing of this region of the genomic clone and comparison to the cDNA clone will identify the translation initiation ATG codon, and if the cDNA is full-length, will give an indication of the transcription start site, upstream of which lies the gene promoter. Important promoter sequence elements may

lie in excess of around 2 kbp of the transcription start site. Therefore, a genomic fragment of preferably 1 - 5 kbp is isolated for initial testing of promoter activity. For example, in the case of U1A and U2B", a monoclonal antibody (mAb), 4G3, raised against a b-galactosidase-human U2B" protein, was used as a probe to screen a potato cDNA expression library and a full-length cDNA clone was isolated (Simpson *et al*. 1991). Screening of a potato genomic library using this cDNA clone resulted in a genomic clone containing part of the gene and around 15 kbp upstream sequences which were used to clone the promoter in subsequent experiments. The extreme similarity between U1A and U2B" enabled us to isolate a full-length genomic clone of the U1A gene. Since the latter clone contained around 7 kbp of upstream sequences, it was possible to clone the promoter from the upstream region.

The use of spliceosomal protein gene promoters to drive expression of DNA sequences in transgenic plants or animals has the advantage that the expression is constitutive, expressed in all cell and tissue types and uses naturally occurring plant or animal nucleic acid sequences which are not derived from infectious agents.

While further modifications and improvements may be made without departing from the scope of this invention, the following is a description of one or more examples of the invention, with reference to the accompanying drawings, in which:-

Fig. 1 shows a schematic diagram of a potato U2B" genomic clone present in the cells of culture No SCRI/JB/2;

Fig. 2 shows a schematic diagram of a potato U1A

genomic clone present in the cells of culture No SCRI/JB/1;

Fig. 3 shows a schematic diagram of a maize PRP8 genomic clone present in the cells of culture No SCRI/JB/3;

Fig. 4 shows a schematic diagram of a UIA/GUS expression cassette;

Fig. 5 shows a schematic diagram of a U2B"/GUS expression cassette;

Fig. 6 shows a graph of fluorescence against time plotted from data collected from a fluorometric assessment of GUS gene expression controlled by potato spliceosomal protein gene promoters;

Fig. 7 shows a schematic diagram of a UIA/GUS expression cassette;

Fig. 8 shows a schematic diagram of U2B"/GUS expression cassette;

Fig. 9 shows a schematic diagram of the genomic organisation of the potato U1A gene; and

Fig. 10 shows a schematic diagram of the genomic organisation of the potato U2B" gene.

Example 1

Genes for two plant spliceosomal proteins from the dicotyledonous plant, potato (*Solanum tuberosum*) were isolated encoding the spliceosomal proteins U1A and U2B".

The promoter of a potato U2B" gene was isolated from a potato genomic library in λ EMBL 3 by conventional methods. The library was screened with a potato U2B" cDNA clone (Simpson et al., 1991) using standard procedures (Sambrook et al., 1989). The potato U2B" genomic clone was plaque-purified, DNA prepared, fragments subcloned into plasmid vectors and DNA sequenced by standard procedures (Sambrook et al.

1989).

The genomic clone contained an insert of 15 kilobase pairs (kbp) (Fig. 1) from which relevant sub-fragments were cloned into plasmid vectors (Fig. 1). The clone only contained a fragment of coding region of U2B" (100 bp), a fragment of the first intron in U2B" (100 bp) and approximately 15 kbp of upstream sequences containing the U2B" promoter.

Example 2

The promoter of a potato U1A gene was isolated from a potato genomic library in λ EMBL 4. The library was screened with a potato U2B" cDNA clone (Simpson et al., 1991) because U1A and U2B" are closely related, and a genomic clone was obtained by known methods. The genomic clone was characterised by state-of-the-art methodologies as described by Sambrook et al., (1989). The clone contained an insert of 15 kbp (Fig. 2) from which relevant sub-fragments were cloned into plasmid vectors by standard techniques (Sambrook et al., 1989). The clone contained the whole of the U1A coding sequence on five exons, four introns and 7 and 2 kbp of 5′ and 3′ flanking sequence respectively.

The promoter regions can be linked to marker genes such as bacterial β-glucuronidase (Jefferson, 1987) by standard molecular techniques (Sambrook et al., 1989). Promoter constructs can be analyzed by introduction into plant cells by known methodology such as chemical or electrical transfection, microinjection, biolistics or *Agrobacterium*-mediated or other vector-mediated transformation (see Shaw, 1988). Transgenic plants containing the construct can be analyzed by detecting the presence of GUS enzyme and thereby its expression using known methods of histochemical staining.

(Jefferson, 1987). Levels of expression of marker genes driven by the promoters in either stably transformed plants or transiently transformed plant cells or protoplasts can be assessed by comparison with levels of endogenous gene expression and of marker gene expression driven by the Cauliflower Mosaic Virus 35S RNA promoter. This analysis can use known, state-of-the-art methodologies to detect RNA transcripts (Sambrook et al., 1989; Simpson et al., 1992) and to detect production of enzyme from marker genes, for example, for GUS marker gene activity.

The promoter region can be incorporated into plasmid vectors designed for general use in construct production in *E.coli*, and for use in stable, *Agrobacterium*-mediated transformation and in transient transformation or stable, physical transformation methods. DNA sequences to be expressed in the transgenic plant can be inserted behind the promoter regions as is currently commonly performed using the CaMV 35S RNA promoter (see Shaw, 1988) prior to introduction into plant cells or production of transgenic plants.

Example 3

A gene for a plant spliceosomal protein was cloned from the monocotyledonous plant, maize (*Zea mays* L.), which encodes the spliceosomal protein PRP8 (Jackson et al., 1988) by virtue of sequence homology to PRP8 of yeast.

The genomic clone of maize PRP8 was isolated from a maize genomic library constructed in λ EMBL 4 by conventional methods. The library was screened with a fragment of maize PRP8 generated by state-of-the-art methodologies, such as polymerase chain reaction (PCR) amplification using oligonucleotide sequences designed

from the yeast and *Caenorhabditis elegans* PRP8 DNA sequences, cloning and sequencing. The maize PRP8 genomic clone was plaque purified, DNA prepared and fragments subcloned into plasmid vectors by standard procedures.

The PRP8 promoter region can be incorporated into plasmid vectors as previously described.

Example 4

**Promoter/GUS Constructions**

Two expression cassettes containing the β-glucuronidase (GUS) gene driven by the potato U1A and U2B" splicesomal protein gene promoters respectively were constructed. Both cassettes contained the nopaline synthase poly A (NOS-ter) downstream to the GUS gene. The construction of these expression cassettes was achieved by replacing the CaMV 35S promoter in the vector pBI221 by upstream sequences of the potato U1A and U2B" spliceosomal protein genes. Genomic clones described above were used.

**U1A**

A 4.5 kb *PstI* fragment containing about 2.5 kb upstream to the U1A gene, the first and the second exons, the first intron and part of the second intron was used to clone the U1A promoter. In order to clone the promoter, a site specific mutation was generated which changed the sequence ATGGCG at the translation start site into TCTAGA to introduce an *XbaI* restriction site. Subsequently, the entire 2.5 kb upstream region was cloned into pBI221 after eliminating the 35S promoter using the restriction enzymes *PstI* and *XbaI* (fig. 4).

**U2B"**

An *EcoRI/SalI* fragment containing about 2 kb upstream to the U2B" gene, the entire first exon, and part of the first intron was used to clone the U2B" promoter. Appropriate restriction sites were introduced into the upstream region using PCR amplification. The first PCR primer started a few nucleotides downstream of the *EcoRI* site and contained additional nucleotides in its 5′ end to provide *PstI* and *BglII* sites. The second primer started two nucleotides upstream of the ATG start codon and contained additional nucleotides providing a site for cleavage by *BamHI* at its 5′ end. After PCR amplification and digestion with *BamHI* and *PstI*, the 2 kb upstream region was cloned into pBI221 replacing the 35S promoter (fig. 5).

**Protoplast Isolation**

The constructed cassettes were tested for promoter activity in transient gene expression assays using tobacco protoplasts. Protoplasts were prepared from young, fully expanded tobacco leaves. Leaves were placed in 7% Domestos for 10 minutes, washed with sterile tap water, dried, and peeled to remove the lower epidermis. Peeled leaf pieces were placed onto 15 ml enzyme solution in a sterile Petri dish [enzyme solution: 1 mg/ml cellulase, 0.5 mg/ml driselase, 0.2 mg/ml macerase, suspended in TO$^-$ (see appendix)] and incubated overnight in the dark at 25°C. Protoplasts were transferred into two sterile 10 ml tubes through sterile sieves and spun at 400 rpm for 5 minutes. After resuspending each protoplast pellet in 10 ml TO$^-$, 5 ml aliquots were each layered onto 2.5 ml 16% sucrose and spun at 800 rpm for 5 minutes. The purified protoplasts were resuspended in 10 ml TO$^-$, spun at 400 rpm for 5 minutes, and recollected in 5 ml TO$^-$.

**Protoplast Transfection**

Transfection of the protoplasts with the U1A/GUS and U2B"/GUS constructs was achieved using plasmid DNA purified by Qiagen (Trade Mark) columns. For each experiment 200 µl of the purified protoplasts were transfected with 30 µg DNA dissolved in 20 µl distilled water. After dropwise addition of the DNA and careful homogenization of protoplasts, 200 ml of PEG solution was added. [PEG solution: 25% PEG 8000, 0.1M $Ca(NO_3)_2$, 0.45M Mannitol, 10 mM MES: 2-(N-Morpholino) ethanesulfonic acid] was added dropwise. The transfected protoplasts were then incubated at ambient temperature for 20 minutes; subsequently, 4 ml calcium nitrate was carefully added [0.275 M $Ca(NO_3)_2$]. After an incubation period of 20 minutes at ambient temperature protoplasts were spun at 400 rpm for 3 minutes, suspended, and resuspended in 5 ml $TO^+$. The transfected protoplasts were then incubated in sealed Petri dishes in the light at 25°C for 48 hours.

**GUS Assays**

Samples were collected by spinning the protoplasts at 400 rpm for 3 minutes, resuspended, transferred into Eppendorf tubes (1 ml each), and spun again at 1000 rpm for 2 minutes in a cold microcentrifuge (4°C). The pellet was then resuspended in 200 µl extraction and reaction buffer (50 mM $NaPO_4$ pH 7.0, 10 mM EDTA, 0.1% Triton X-100, 10 mM β-mercaptoethanol), spun at 1300 rpm for 3-5 minutes, and the supernatant was stored in clean Eppendorf tubes. 50 µl sample extract was added to 220 µl extraction and reaction buffer in an eppendorf tube and preincubated at 37°C for 5 minutes. At 15 sec intervals, 30 µl 1 mM MUG (4-methyl umbelliferyl glucuronide) substrate was added to each tube giving time to stop reactions at accurate intervals. At defined time points, 100 µl samples were

taken and added to 2.9 ml stop solution (0.2 M $Na_2CO_3$) in a clean cuvette. At this stage, GUS activity could be inspected visually by transillumination using a long wave UV light box. Quantitative measurements of fluorescence were made using a fluorometer.

Protoplast transfections and GUS assays have been performed to investigate the expression of the GUS gene driven by splicesomal protein gene promoters using the above described constructs. In all experiments, the vector pBI221 which contains the GUS gene driven by the CaMV 35S promoter and terminated by the nopaline synthase poly A was used as a positive control. Negative controls were a: extracts from protoplasts transfected with water instead of plasmid DNA; and b: MUG substrate added to the extraction and reaction buffer without the addition of any extracts. Reactions were carried out for 30 minutes, 1 hour, 2 hours, 3 hours, 4 hours, 5 hours and 20 hours. In order to assess the activity of the tested promoters in comparison to the 35S promoter, fluorometer readings were set at 100 for pBI221 at each time point so that the readings obtained for U1A and U2B" indicated the GUS activity in each case relative to that obtained for the 35S promoter. The results which are illustrated in fig 6 show that the U2B" promoter has about 70% of the activity of the 35S promoter whereas the U1A promoter gave rise to 10-15% of the activity observed for the 35S promoter.

Figs. 9 and 10 show the genomic organisation of the original genomic clones made of potato genes U1A and U2B". The promoter of the U1A gene is believed to lie in the region between the ATG initiation codon and the *Pst*I site 2.5 kb upstream of the coding sequence initiating at the ATG start codon. The promoter of the

U2B" gene is believed to lie in the region between the ATG coding region and the *Eco*RI site 2.0 kb upstream of the coding sequence initiated by the ATG start codon.

Figs. 9 and 10 indicate regions of the genomic clones containing the promoters of U1A and U2B" which have been sequenced. The sequences of those regions are shown in the sequence listing.

The partial U1A promoter sequence is shown on the sequence listing as SEQ:ID:No1. Sequence data is presented for the 3′ end of the region between the ATG codon and a *Pst*1 site 2.5kb upstream. About 1.7kb of the 5′ end of this region has not been sequenced. The partial sequence of the U2B" genomic clone is shown in the sequence listing as three separate sequences since incomplete sequence data is available for this promoter. The first of the U2B" sequences shown is SEQ:ID:No2 which corresponds to the region from the 5′ end TCT to a TAA 119 bases downstream. After the 3′ end of SEQ:ID:No2 there is a region of approximately 900 bases which have not been sequenced. After the 3′ end of the 900bp unidentified region, the next U2B" sequence portion available is SEQ:ID:No3 which runs from a TCT approximately 900 bases downstream of the end of SEQ:ID:No2 to an ATC a further 139 bases downstream. After the 3′ end of SEQ:ID:No3 there is a region of 11 bases which have not been identified. This unidentified region is followed by the last sequenced portion of the U2B" clone which is designated as SEQ:ID:No4 which starts at ATA at its 5′ end and ends at the ATG codon (bases 729-731) initiating transcription of the coding region of U2B" (not shown).

Spliceosomal protein gene promoters (or other expression control polynucleotides) have the advantages

that (a) spliceosomal proteins are absolutely required and thus spliceosomal protein gene promoters are likely to be active in every cell and tissue type; (b) they are not derived from infectious agents which overcomes objections to the use of such sequences due to potential recombination; and (c) different genes or DNA sequences are likely to be expressed at different levels reflecting the relative abundance of the different spliceosomal proteins.

Modifications and improvements may incorporated without departing from the scope of the invention. For example, an expression control polynucleotide in accordance with the invention may be operably linked to a second polynucleotide which has some function other than coding for a polypeptide. One such example might be to operably link an expression control polynucleotide to a gene encoding a ribozyme or anti-sense RNA. Indeed it will be realised by the skilled man that the function of the gene under the control of the expression control polynucleotide of the invention is not important, and that the expression of many diverse genes can be controlled.

References

The following references are incorporated herein by reference.

Jackson, S.P., Lossky, M & Beggs, J.D. (1988). Mol Cell Biol. 8, 1067-1075.

Jefferson, R.A. (1987). Assaying chimaeric genes in plants: the GUS gene fusion system. Plant Molecular Biology Reporter 5, 387-405.

Simpson, G.G., Vaux, P., Clark, G., Waugh, R., Beggs, J.D. & Brown, J.W.S. (1991). Evolutionary conservation of the spliceosomal protein, U2B". Nucleic Acids Research 19, 5213-5217.

Simpson, C.G., Sinibaldi, R. & Brown, J.W.S. (1992). Rapid analysis of plant gene expression by a novel reverse transcriptase-PCR method. Plant Journal 2, 835-836.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shaw, C.H. (1988). Plant molecular biology: a practical approach. IRL Press.

Appendix 1

TO-:

4 ml solution 1
200 ml solution 3
200 ml solution 4
200 ml NAA (3 mg/ml)
200 ml BAP (1 mg/ml)
16 g Mannitol
pH 5.5 adjusted with NaOH
filtersterilised +200 ml
cefotaxime (100 mg/ml)

TO+:

4 ml solution 1
4 ml solution 2
200 ml solution 3
200 ml solution 4
200 ml NAA (3 mg/ml)
200 ml BAP (1mg/ml)
16 g Mannitol
4 g Sucrose
40 ml Tween 20
pH 5.5 adjusted with NaOH
filtersterilised +200 ml
cefotaxime (100 mg/ml)

**Solution 1:**

10.30 mM $NH_4NO_3$
9.40 mM $KNO_3$
1.50 mM $CaCl_2.2H_2O$
0.75 mM $MgSO_4.7H_2O$
0.62 mM $KH_2PO_4$

**Solution 2:**

100 $\mu$M $FeSO_4$
100 $\mu$M $Na_2$ EDTA

**Solution 3:**

16.00 $\mu$M $H_3BO_3$
0.60 $\mu$M $MnSO_4.H_2O$
3.50 $\mu$M $ZnSO_4.7H_2O$
0.12 $\mu$M $CuSO_4.5H_2O$
0.22 $\mu$M $AlCl_3$
0.13 $\mu$M $NiCl_3.6H_2O$
0.06 $\mu$M KI

**Solution 4:**

555.00 $\mu$M Inostol
3.00 $\mu$M Thiamine
5.00 $\mu$M Pyridoxine
8.00 $\mu$M Niacin
2.00 $\mu$M Pantothenate
0.04 $\mu$M Biotine
+ 1 mg/$\mu$l NaOH

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Scottish Crop Research Institute
(B) STREET: Invergowrie
(C) CITY: Dundee
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): DD2 5DA
(G) TELEPHONE: +44 1382 562731
(H) TELEFAX: +44 1382 561442

(ii) TITLE OF INVENTION: Expression control polynucleotides derived from spliceosomal protein gene promoters

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 784 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TTAGAATTAG AATCCCCATT TTTAAGAATA ATCCTAGATA ATTTTCTTAA ACATGACAAT 60

TGATACCCAC AATTAATTAC TATTACATAA ATTTTTACCT AAATTAGATA TAACTTTCAA 120

TTTCAAAAAT TAAAACCCAA AAAAATTGAA CGACAATACG AGAGGGGATC AAACATAGGC 180

GAGCAATTAG AGAAATTGAC GGGTAGACAT CAACAAACCA TCAAGAATTT AAAAGCGGAA 240

AGAGAAAAAA ATACACTATG GACGAATATT TTTATAGAAT TCAATATGTA AAACTAATAA 300

ACAAGAAAGT AAATCATCTT TTATTCAAAG TAATGAAGAA GAAGAATTGA ATAAATATTT 360

ACATAATCAA TAAAAAAAAC TCATTCAAAA GAATCGTGTG TATGGGAAAG AAGAAGAAAA 420

AAAAGGCAGA AAAAAACCAC TTCCCAATAA AAAAGGACAT CATGCTGCCA CCTCCTAAAA 480

TTATTTAATT TAATTAAAAA AAAAACTTCC CAACACGTGG GCTACTAATT GCAAAATTTA 540

ATTTTTAAAA AGCTTTTTTT GTCAAGAAAA TAAAAGATGG CTATATGTTG CCAATTAGTA 600

AAATGGGATG TCATGCTGTG TCATTTTTTC TTGAGTTGTT AAGGGCTCAA AGCCCAATTG 660

TTTATCCAGC CCAAGCCCAA ATCGGAGCCC TATTCGTGCC CAAAAATTTC TGGAGAAATT 720

AACGACAACT GAAGTTTCTA CCTCACCGGC GAAAGTTGCA GCTAGGCGTA GACGAGGAGC 780

CATG 784

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TCTCTCCAGC TCTTCCCTCC TAAAACAACC ATTTTATGAG TACAGACACA AACCAGCTTA 60

GCAACCAGTA AATCCAAAAC TTTAATTCCA CGTGTAAGCG CTAACACTTC ACCCACTAA 119

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TCTTTACACT ATTTTCAAAT AACGATGAGA CCTGTAATAA TGTAATAACT TGAAAATAGA 60

ACAATAACTC ATTCAGTACA ACAAATAAAA TCATACTAAT GTATATTTTT AAAAACAATT 120

TAACTCTATT TAATTAATC 139

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 731 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATATCTAAAA AAATGTGATT GGAGCACTCA TTGACCACAC GAATGAATCT ACAATGTAGA 60

TCATTCTATC TAGACAAATA ATGCAAAAAC TAAAAGATAA AAGTAATATT ATACTACAAT 120

CTGTTAAATG ATATCAATAT TACAAAAGTT CTCTACGATG TCAATACATA TATTAAAATC 180

TATTTGATTA ATCAGAAACA TATCATGTGT GAATTTTTTT AATTAAAGAT CCCTTTAATC 240

ATCTGAATCA ACCTTGGCTG GTCTCACATC TTTCCACCCT CTACTCGGTC TTCACTTTCT 300

CTTAAACTAG GGAAGAACAA CATGATATTA GCTTAGATTA ATTAAACAAG CTCATCAAAA 360

CTACCATCCA ATTTAAGCCA ATAATGTTTA AATAAAACAA AAAACAACGT ACTCATTTTT 420

TCATAACAAG AAGTTAAAAT TTATATGAAT CCTTACTCCA AAAAAGAAGA AAAATTACAA 480

TATCAATATA TATAACTACT CTATTTGGTT AGTCAACAAA ATGTTAGTAT ATGTATTGCA 540

AGTTCGCAAC ACCCGCTTGG GCCTTGACCA CATATTTATA TGGGCCGGTT GTCAATTTAA 600

GCCCACTTTG TGTTCGTTCG CCTTTCTTGT AGCTCCAAAC TCTTGGAAAT TTGTCGAGCA 660

CATTCAGAAA TCACAGAGAA GAGCAAGTGA ATATACATAC AGATAGAGAA AAGCTGCTCT 720

GCTCGGTAAT G 731

Claims

1 An expression control polynucleotide at least partially derived from a spliceosomal protein gene promoter.

2 An expression control polynucleotide according to Claim 1 derived from a spliceosomal protein gene promoter of a plant.

3 A recombinant polynucleotide comprising an expression control polynucleotide according to Claim 1 or Claim 2 operably linked to a second polynucleotide.

4 A recombinant polynucleotide according to Claim 3, wherein the second polynucleotide encodes a polypeptide.

5 A recombinant polynucleotide according to Claim 3, wherein the second polynucleotide encodes a ribozyme.

6 A recombinant polynucleotide according to Claim 3, wherein the second polynucleotide encodes anti-sense RNA.

7 A recombinant vector containing an expression control polynucleotide according to Claim 1 or Claim 2, or containing a recombinant polynucleotide according to any one of Claims 3 to 6.

8 A method for producing a recombinant vector, said method comprising ligating an expression control polynucleotide according to Claim 1 or Claim 2 to

a vector.

9 A transformed host cell containing a recombinant polynucleotide according to any one of Claims 3 to 6, or a recombinant vector as claimed in Claim 7.

10 A transgenic organism and/or the progeny and/or seeds thereof, containing a recombinant polynucleotide according to any one of Claims 3 to 6, or a recombinant vector according to Claim 7.

11 A method for controlling the expression of a polypeptide from a polynucleotide encoding the polypeptide, said method comprising operably linking said polynucleotide to an expression control polynucleotide according to Claim 1 or Claim 2.

12 A spliceosomal protein gene promoter deposited as any one of NCTC 12864, NCTC 12865 and NCTC 12866.

13 A recombinant polynucleotide having a spliceosomal protein gene promoter according to Claim 12.

14 A recombinant vector having a spliceosomal protein gene promoter according to Claim 12.

15 A transformed host cell containing a spliceosomal protein gene promoter as claimed in Claim 12, a recombinant polynucleotide as claimed in Claim 13, and/or a recombinant vector as claimed in Claim 14.

16 A spliceosomal protein gene promoter comprising a region of a plasmid in NCTC 12864 between an ATG codon at the start of the coding sequence and a

*Pst*I site 2.5 kb upstream of the ATG codon.

17 A spliceosomal protein gene promoter comprising a region of a plasmid in NCTC 12865 between an ATG codon at the start of the coding sequence and an *Eco*RI site 2 kb upstream of the ATG codon.

18 A polynucleotide having any one of the sequences SEQ:ID:No1, SEQ:ID:No2, SEQ:ID:No3, SEQ:ID:No4 and functional equivalents thereof.

FIG. 1

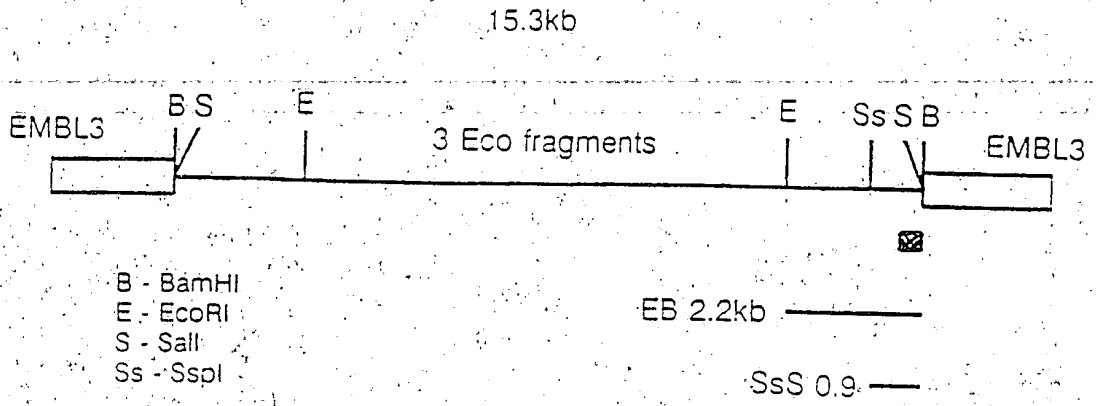

15.3kb
EMBL3
B S
E
3 Eco fragments
E
Ss S B
EMBL3
B - BamHI
E - EcoRI
S - SalI
Ss - SspI
EB 2.2kb
SsS 0.9

FIG. 2

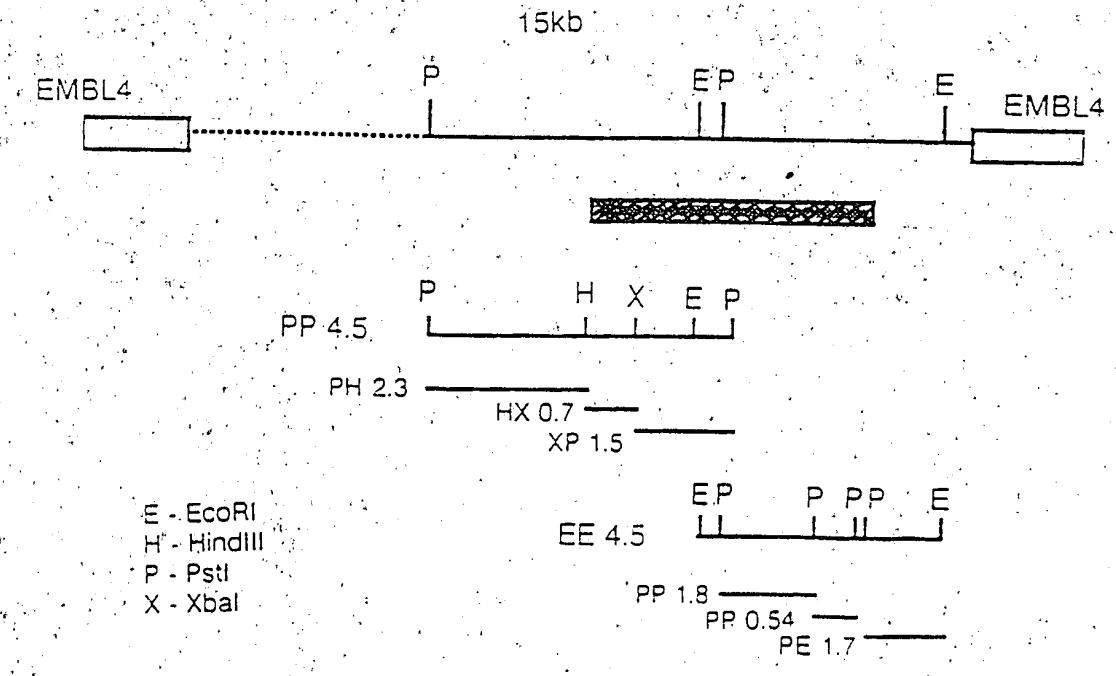

15kb
EMBL4
P
E P
E
EMBL4
P H X E P
PP 4.5
PH 2.3
HX 0.7
XP 1.5
E - EcoRI
H - HindIII
P - PstI
X - XbaI
E P P P P E
EE 4.5
PP 1.8
PP 0.54
PE 1.7

FIG. 3

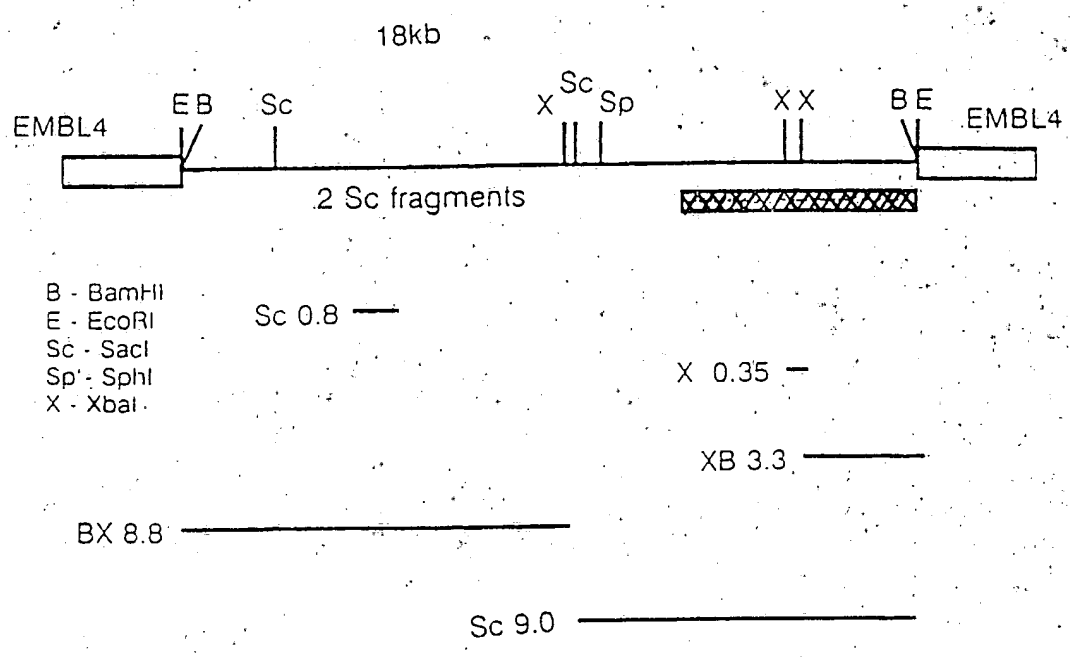

18kb
EMBL4
E B
Sc
X
Sc
Sp
X X
B E
EMBL4
2 Sc fragments
B - BamHI
E - EcoRI
Sc - SacI
Sp - SphI
X - XbaI
Sc 0.8
X 0.35
XB 3.3
BX 8.8
Sc 9.0

FIG. 4

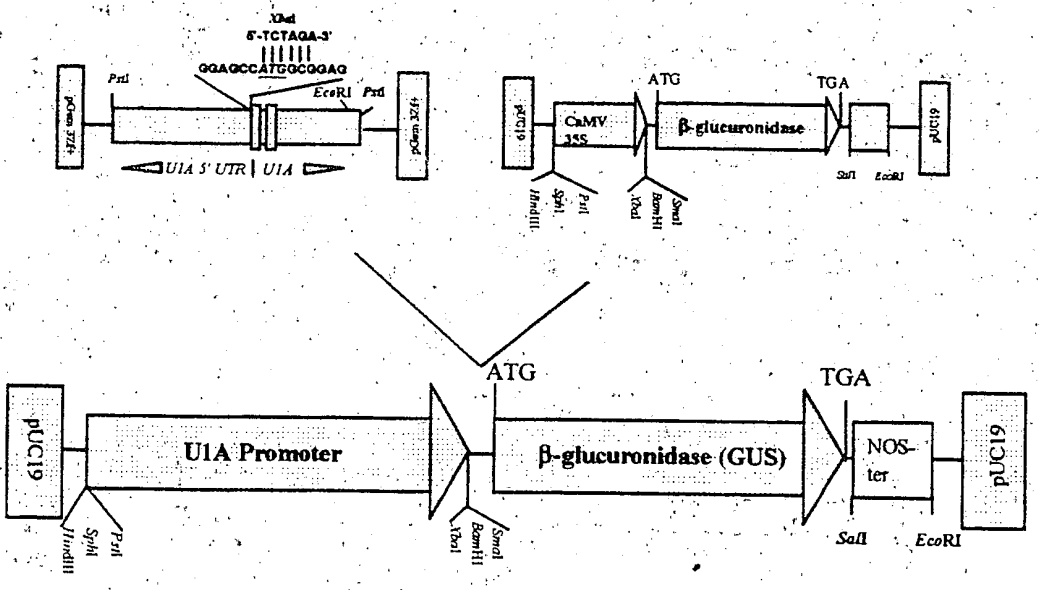

XbaI
5'-TCTAGA-3'
GGAGCCATGGCGGAG
PstI
EcoRI
PstI
pUC19
U1A 5' UTR
U1A
pUC19
CaMV 35S
ATG
β-glucuronidase
TGA
pUC19
SalI
EcoRI
HindIII
SphI
PstI
XbaI
BamHI
SmaI
pUC19
U1A Promoter
ATG
β-glucuronidase (GUS)
TGA
NOS-ter
pUC19
SalI
EcoRI

FIG. 5

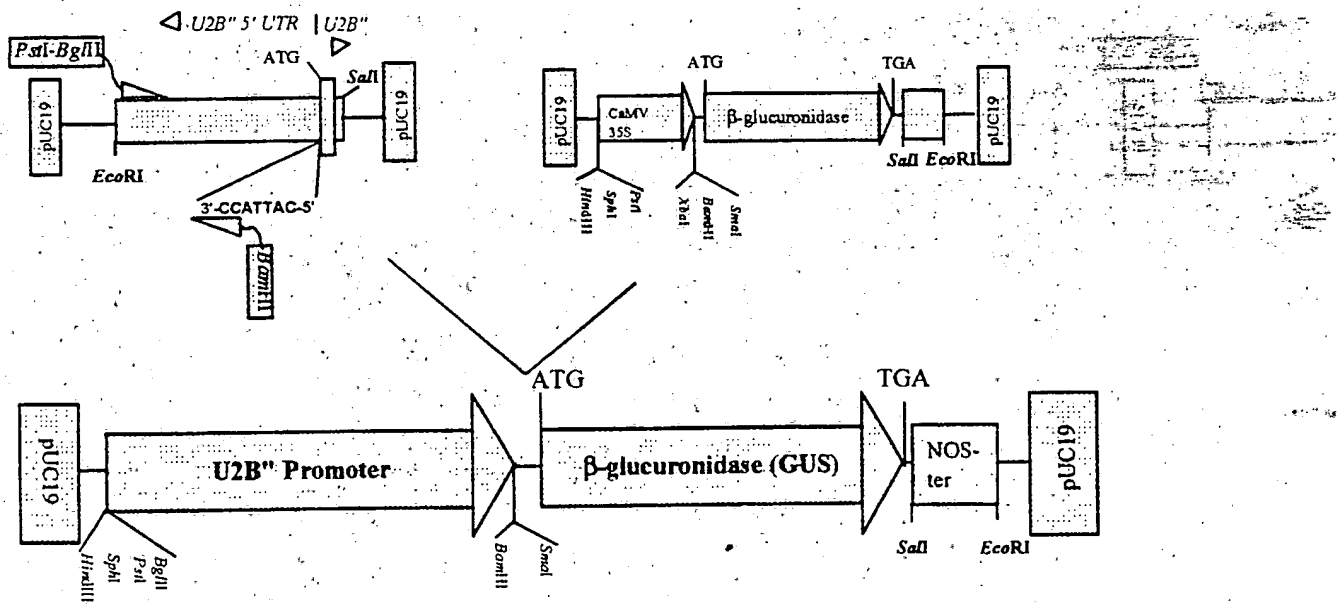

U2B" 5' UTR
U2B"
PstI-BglII
ATG
SalI
pUC19
EcoRI
3'-CCATTAC-5'
BamHI
CaMV 35S
β-glucuronidase
TGA
SalI EcoRI
HindIII
SphI
PstI
XbaI
BamHI
SmaI
U2B" Promoter
β-glucuronidase (GUS)
NOS-ter
BglII

FIG. 6

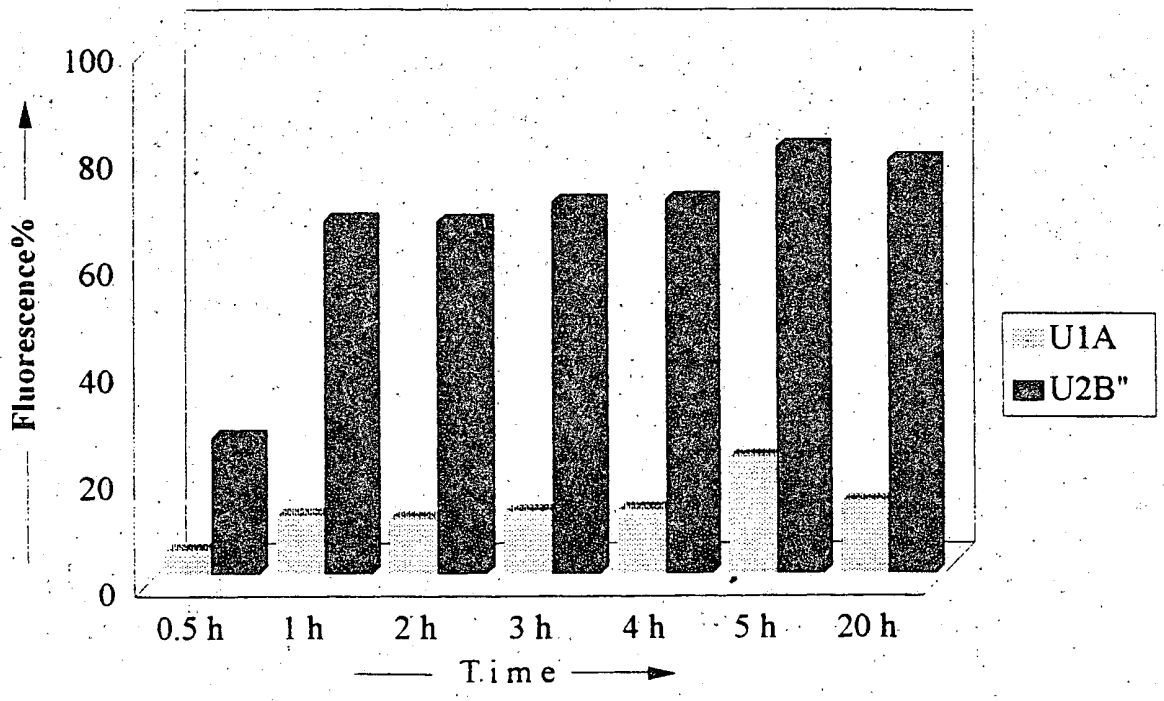

100
80
60
40
20
0
Fluorescence%
0.5 h
1 h
2 h
3 h
4 h
5 h
20 h
Time
U1A
U2B"

FIG. 7

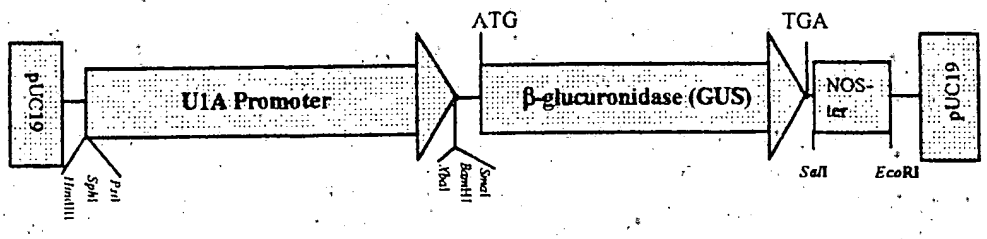

ATG
TGA
pUC19
U1A Promoter
β-glucuronidase (GUS)
NOS-ter
pUC19
HindIII
SphI
PstI
XbaI
BamHI
SmaI
SalI
EcoRI

FIG. 8

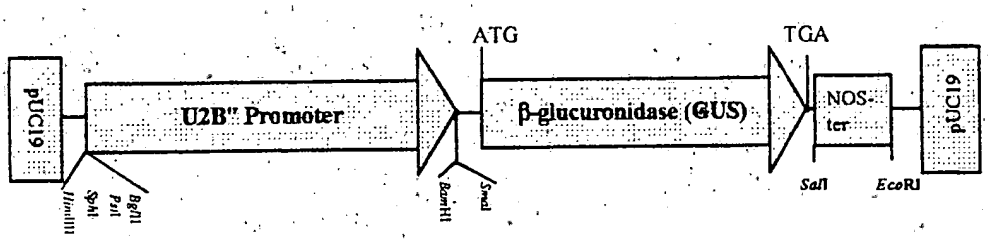

ATG
TGA
pUC19
U2B" Promoter
β-glucuronidase (GUS)
NOS-ter
pUC19
HindIII
SphI
PstI
BglII
BamHI
SmaI
SalI
EcoRI

FIG. 9

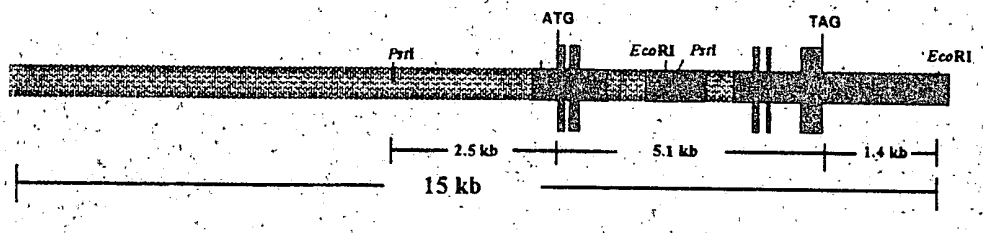
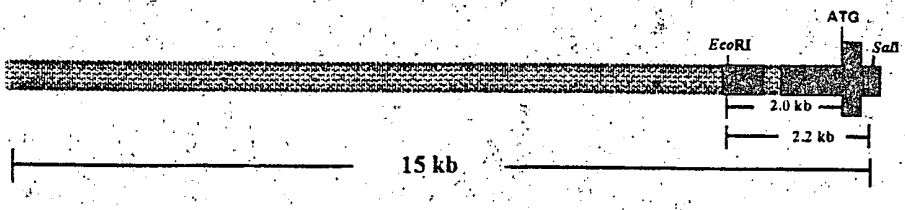
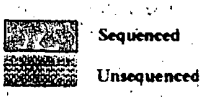

ATG
TAG
PstI
EcoRI
PstI
EcoRI
2.5 kb
5.1 kb
1.4 kb
15 kb

FIG. 10


ATG
EcoRI
SalI
2.0 kb
2.2 kb
15 kb


Sequenced
Unsequenced

# INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/GB 95/01443

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N15/29 A01H5/00 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | MOL. CELL. BIOL., vol. 8, no. 3, 1988 pages 1067-1075, JACKSON, S.P., ET AL. 'CLONING OF THE RNA8 GENE OF SACCHAROMYCES CERVISIAE, DETECTION OF THE RNA8 PROTEIN, AND DEMONSTRATION THAT IT IS ESSENTIAL FOR NUCLEAR PRE-MRNA SPLICING' see page 1068, right column, last paragraph --- | 1 |
| X | NUCLEIC ACIDS RESEARCH, vol. 22, no. 11, 11 June 1994 pages 1996-2002, BIAMONTI, G., ET AL. 'TWO HOMOLOGOUS GENES, ORIGINATED BY DUPLICATION, ENCODE THE HUMAN HNRNP PROTEINS A2 AND A1' see page 2001, left column, paragraph 1 --- -/-- | 1,3,4, 7-11 |

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 2 November 1995 | 22. 11. 95 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016 | Maddox, A |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 95/01443

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | J. MOL. BIOL.,<br>vol. 230, 5 March 1993<br>pages 77-89,<br>BIAMONTI, G., ET AL. 'HUMAN HNRNP PROTEIN A1 GENE EXPRESSION STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE PROMOTER'<br>see the whole document<br>--- | 1,3,4,<br>7-11 |
| P,X | JOURNAL OF EXPERIMENTAL BOTANY,<br>vol. 46, no. SUPPL, 1995<br>page 38<br>HAMILTON, J.I., ET AL. 'CHARACTERISATION OF A GENE ENCODING THE SPLICEOSOMAL PROTEIN PRP8 FROM MAIZE'<br>see abstract<br>& ANNUAL MEETING OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY, ST. ANDREWS, SCOTLAND, UK, APRIL 3-7, 1995.,<br>--- | 1,2 |
| A | PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON B BIOLOGICAL SCIENCES 342 (1301). 1993. 217-224.,<br>BROWN J W S , ET AL. ' Plant pre-mRNA splicing and splicing components.'<br>see page 219, left column<br>see page 222, right column<br>--- | 1-18 |
| A | NUCLEIC ACIDS RESEARCH,<br>vol. 19, no. 19, 1991<br>pages 5213-5217,<br>SIMPSON, G.G., ET AL. 'EVOLUTIONARY CONSERVATION OF THE SPLICEOSOMAL PROTEIN, U2B"'<br>cited in the application<br>see the whole document<br>--- | 1-18 |
| A | FEBS LETTERS,<br>vol. 318, no. 1, February 1993<br>pages 4-6,<br>KULESZA, H., ET AL. 'DETECTION OF A PLANT PROTEIN ANALAGOUS TO THE YEAST SPLICEOSOMAL PROTEIN, PRP8'<br>see the whole document<br>--- | 1-18 |
| A | NUCLEIC ACIDS RESEARCH,<br>vol. 18, no. 8, 1990<br>pages 1941-1949,<br>KISS, T., ET AL. 'MOLECULAR ANALYSIS OF A U3 RNA GENE LOCUS IN TOMATO: TRANSCRIPTION SIGNALS, THE CODING REGION, EXPRESSION IN TRANSGENIC TOBACCO PLANTS AND TANDEMLY REPEATED PSEUDOGENES'<br>see page 1946, right column<br>--- | 1-18 |

-/--

1

# INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/GB 95/01443

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ' | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | NUCLEIC ACIDS RESEARCH, vol. 21, no. 7, 1993 pages 1533-1540, EDOH, D., ET AL. 'ACTIVITY OF U-SNRNA GENES WITH MODIFIED PLACEMENT OF PROMOTER ELEMENTS IN TRANSFECTED PROTOPLASTS AND STABLY TRANSFORMED TOBACCO' see the whole document | 1-18 |